5 **METHOD AND APPARATUS FOR GENERATING METADATA FOR A DOCUMENT**

ABSTRACT

A method and system of generating metadata for a document so that the document may
10 be identified by a subsequent search. A conceptual model is generated for the document,
wherein the conceptual model indicates one or more concepts that are recognized in the
document. A concept is defined by a plurality of features, each feature being associated with a
feature weight. By referencing the conceptual model, one or more auto-attributes may be
assigned to the document. Also, by referencing the conceptual model, the document may be
15 categorized to one or more categories of a categorization taxonomy by assigning one or more
auto-categories. The generated metadata, including the conceptual model, the one or more auto-
attributes, and the one or more auto-categories, may be stored in a memory so that the
subsequent search may identify the document by examining the generated metadata.